

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-324095
 (43)Date of publication of application : 08.11.2002

(51)Int.Cl. G06F 17/50
 C07B 61/00
 G06F 17/15

(21)Application number : 2001-129216 (71)Applicant : NIPPON ZOKI PHARMACEUT CO LTD

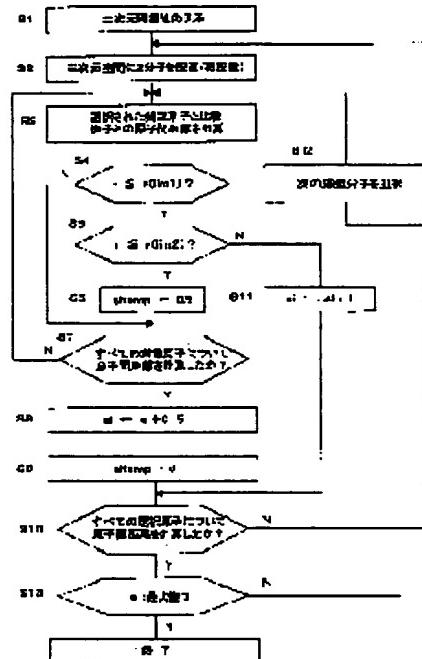
(22)Date of filing : 26.04.2001 (72)Inventor : KOTANI TAKAYUKI

(54) METHOD AND PROGRAM FOR EVALUATION OF MOLECULAR SIMILARITY

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a method and program capable of evaluating similarity between two molecules with calculating time and amount less than those of in methods of Merer and Richards.

SOLUTION: A method for evaluation of molecular similarity comprises a step (a) of acquiring an upper threshold and lower threshold from any of a proper value of an atom included in a first molecule, a proper value of an atom of a second molecule evaluating interrelationship with the atom or other value acquired from those values, a step (b) of calculating interrelationship between atoms of the first and second molecules by using the thresholds and a step (c) of evaluating the similarity between the first and second molecules based on the interrelationship acquired in the step (b).



(19)日本国特許庁(JP)

(12)公開特許公報 (A)

(11)特許出願公開番号

特開2002-324095

(P2002-324095A)

(43)公開日 平成14年11月8日(2002.11.8)

(51)Int.C1.⁷
G 06 F 17/50
C 07 B 61/00
G 06 F 17/15

識別記号
6 3 8

F I
G 06 F 17/50
C 07 B 61/00
G 06 F 17/15

テ-マコード(参考)
4H006
Z 5B046
5B056

審査請求 未請求 請求項の数 2 2 O L

(全14頁)

(21)出願番号 特願2001-129216(P2001-129216)

(22)出願日 平成13年4月26日(2001.4.26)

(71)出願人 000231796
日本臓器製薬株式会社
大阪府大阪市中央区平野町2丁目1番2号

(72)発明者 小谷 孝行
兵庫県加東郡社町木梨川北山442-1 日本
臓器製薬株式会社生物活性科学研究所内

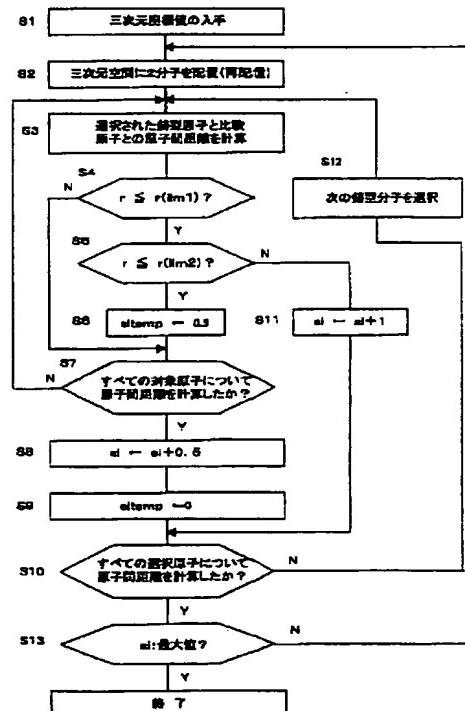
(74)代理人 100062144
弁理士 青山 葦 (外1名)
F ターム(参考) 4H006 AA05 AC90
5B046 AA00 BA05 JA01
5B056 BB23 HH00

(54)【発明の名称】分子類似性評価方法及びプログラム

(57)【要約】

【課題】 Mererら及びRichardsらの方法よりもより少ない計算時間及び計算量で、2つの分子の類似性を評価できる方法とプログラムを提供する。

【解決手段】 分子類似性評価方法は、第1の分子に含まれる原子に固有の値、又は該原子との相関性を評価する第2の分子の原子に固有の値、若しくはそれらの値から求まる別の値のうちのいずれかの値から上限閾値と下限閾値とを求めるステップと、第1の分子の原子と第2の分子の原子との相関性を上限閾値と下限閾値とを用いて計算するステップと、ステップ(b)で得られた相関性をもとに第1の分子と第2の分子の類似性を評価するステップとを含む。



【特許請求の範囲】

【請求項1】 第1の分子と第2の分子との類似性を評価する方法であって、(a) 第1の分子に含まれる原子に固有の値、又は該原子との相関性を評価する第2の分子の原子に固有の値、若しくはそれらの値から求まる別の値のうちのいずれかの値をもとに上限閾値と下限閾値とを求めるステップと、(b) 第1の分子の原子と第2の分子の原子との相関性を上限閾値と下限閾値とを用いて計算するステップと、(c) ステップ(b)で得られた相関性をもとに第1の分子と第2の分子の類似性を評価するステップとを備えたことを特徴とする分子類似性評価方法。

【請求項 2】 第1の分子に含まれる各原子と第2の分子に含まれる各原子との原子間距離をもとに原子間の相関性を求め、この相関性から第1と第2の分子の類似性を評価する方法であつて、

2つの原子間の相関性を求める処理は、(a) 第1の分子の原子及び／又は第2の分子の原子に固有の値を求めるステップと、(b) ステップ(a)で求めた固有の値から上限閾値を求めるステップと、(c) ステップ(a)で求めた固有の値から下限閾値を求めるステップと、(d) 2つの原子間の距離を求めるステップと、(e) ステップ(d)で求めた原子間の距離と、上限閾値と、下限閾値とを用いて、第1の分子の原子と第2の分子の原子との相関性を評価するステップとを含むことを特徴とする分子類似性評価方法。

【請求項3】 三次元座標系に配置された第1の分子と第2の分子との類似性を評価する方法であって、(a)

相関性指数をリセットするステップと、(b) 第2の分子に含まれる複数の原子の中から一つの原子を選択するステップと、(c) 第1の分子に含まれる複数の原子の中から一つの原子を選択するステップと、(d)

ステップ (b) と (c) でそれぞれ選択された原子の双方又は一方に固有の値から上限閾値と下限閾値とを求めるステップと、(e) ステップ (b) と (c) でそれぞれ選択された原子間距離を求めるステップと、

(f) ステップ (d) で求めた上限閾値と下限閾値をもとに、ステップ (e) で求めた原子間距離に応じた相関性指数を求めるステップと、(g) ステップ (f) の終了後、第 1 の分子に含まれる複数の原子の中から未だ選択されていない原子が有るか否か判断するステップと、(h) ステップ (g) で、第 1 の分子に含まれる複数の原子の中から未だ選択されていない原子が有ると判断されると、ステップ (c) に戻り第 1 の分子に含まれる複数の原子の中から未だ選択されていない他の原子を選択し、ステップ (d) ~ (f) を実行して相関性指数を更新するステップと、(i) ステップ (g) で、第 1 の分子に含まれる複数の原子の中から未だ選択されていない原子が無いと判断されると、ステップ (a) に戻り第 2 の分子に含まれる複数の原子の中から未だ選択

されていない他の原子を選択し、ステップ(c)～(h)を実行し、相関性指数を加算するステップとを有することを特徴とする分子類似性評価方法。

【請求項4】 ステップ(e)で求める原子間距離は、直交座標系におけるいずれかの軸に関する距離であることを特徴とする請求項3に記載の分子類似性評価方法

【請求項5】 2つの分子の類似性を評価する方法において、(a) 第1及び第2の分子に含まれる各原子の三次元座標を得るステップと、(b) 第1及び第2の分子を三次元座標系に配置するステップと、(c) 第2の分子に含まれる各原子について第1の分子に含まれる各原子との類似性指数を求めるステップであって、

(c-0) 相関性指数をリセットするサブステップと、(c-1) 第2の分子に含まれる複数の原子の中から一つの原子を選択するサブステップと、(c-2)

第1の分子に含まれる複数の原子の中から一つの原子を選択するサブステップと、(c-3) サブステップ(c-1)と(c-2)でそれぞれ選択された原子の双方又は一方に固有の値から上限閾値と下限閾値とを求めるサブステップと、(c-4) サブステップ(c-1)と(c-2)でそれぞれ選択された原子間距離を求めるサブステップと、(c-5) サブステップ(c-

3) で求めた上限閾値と下限閾値をもとに、サブステップ (c-4) で求めた原子間距離に応じた相関性指数を求めるサブステップと、(c-6) サブステップ (c-5) の終了後、第 1 の分子に含まれる複数の原子の中から未だ選択されていない原子が有るか否か判断するステップと、(c-7) サブステップ (c-6) で、第

1の分子に含まれる複数の原子の中から未だ選択されていない原子が有ると判断されると、サブステップ(c-2)に戻り第1の分子に含まれる複数の原子の中から未だ選択されていない他の原子を選択し、サブステップ

(c-3) ~ (c-6) を実行して相関性指数を更新するステップと、(c-8) サブステップ (c-6) で、第 1 の分子に含まれる複数の原子の中から未だ選択されていない原子が無いと判断されると、ステップ (c-

- 0) に戻り第2の分子に含まれる複数の原子の中から未だ選択されていない他の原子を選択し、ステップ(c-2)～(c-7)を実行し、相関性指数を加算して類似度合を算出する。

似性指数を求めるサブステップとを有するステップと、
(d) ステップ(c)の終了後、三次元座標内で第1の分子と第2の分子との相対位置を変更するステップとを有することを特徴とする分子類似性評価方法。
【請求項6】 サブステップ(c-4)で求める原子間距離は、直交座標系におけるいすれかの軸に関する距離であることを特徴とする請求項5に記載の分子類似性評価方法。

【請求項 7】 ステップ(d)は、類似性指数が高まる方向に向けて、第1の分子と第2の分子との相対位置を変更することを特徴とする請求項5に記載の分子類似性

評価方法。

【請求項8】 原子に固有の値は、該原子のファンデルワールス半径、原子半径、共有半径のいずれかであることを特徴とする請求項1から7のいずれか一に記載の分子類似性評価方法。

【請求項9】 上限閾値は、原子に固有の値に上限係数を乗じて得られる値であることを特徴とする請求項1から7のいずれか一に記載の分子類似性評価方法。

【請求項10】 下限閾値は、原子に固有の値に下限係数を乗じて得られる値であることを特徴とする請求項1から7のいずれか一に記載の分子類似性評価方法。

【請求項11】 上限係数が1.0～1.3の範囲内の値であることを特徴とする請求項9に記載の分子類似性評価方法。

【請求項12】 下限係数が0.5～0.8の範囲内の値であることを特徴とする請求項10に記載の分子類似性評価方法。

【請求項13】 三次元座標系に配置された第1の分子と第2の分子との類似性を評価するプログラムであって、計算機に(a) 相関性指数をリセットする処理と、(b) 第2の分子に含まれる複数の原子の中から一つの原子を選択する処理と、(c) 第1の分子に含まれる複数の原子の中から一つの原子を選択する処理と、(d) 処理(b)と(c)でそれぞれ選択された原子の双方又は一方に固有の値から上限閾値と下限閾値とを求める処理と、(e) 処理(b)と(c)でそれぞれ選択された原子間距離を求める処理と、(f) 処理(d)で求めた上限閾値と下限閾値をもとに、処理(e)で求めた原子間距離に応じた相関性指数を求める処理と、(g) 処理(f)の終了後、処理(c)に戻り第1の分子に含まれる複数の原子の中から未だ選択されていない他の原子を選択し、ここで選択された他の原子について処理(d)～(f)を実行して相関性指数を更新する処理と、(h) 処理(g)の終了後、処理(a)に戻り第2の分子に含まれる複数の原子の中から未だ選択されていない他の原子を選択し、ここで選択された他の原子について処理(c)～(g)を実行し、相関性指数を加算する処理とを実行させることを特徴とする分子類似性評価プログラム。

【請求項14】 処理(e)で求める原子間距離は、直交座標系におけるいずれかの軸に関する距離であることを特徴とする請求項13に記載の分子類似性評価プログラム。

【請求項15】 2つの分子の類似性を評価するプログラムであって、計算機に、(a) 第1及び第2の分子に含まれる各原子の三次元座標を得る処理と、(b) 第1及び第2の分子を三次元座標系に配置する処理と、(c) 第2の分子に含まれる各原子について第1の分子に含まれる各原子との類似性指数を求める処理であつて、(c-0) 相関性指数をリセットするサブ処理

と、(c-1) 第2の分子に含まれる複数の原子の中から一つの原子を選択するサブ処理と、(c-2) 第1の分子に含まれる複数の原子の中から一つの原子を選択するサブ処理と、(c-3) サブ処理(c-1)と(c-2)でそれぞれ選択された原子の双方又は一方に固有の値から上限閾値と下限閾値とを求めるサブ処理と、(c-4) サブ処理(c-1)と(c-2)でそれぞれ選択された原子間距離を求めるサブ処理と、(c-5) サブ処理(c-3)で求めた上限閾値と下限閾

10 閾値をもとに、サブ処理(c-4)で求めた原子間距離に応じた相関性指数を求めるサブ処理と、(c-6) サブ処理(c-5)の終了後、サブ処理(c-2)に戻り第1の分子に含まれる複数の原子の中から未だ選択されていない他の原子を選択し、ここで選択された他の原子についてサブ処理(c-3)～(c-5)を実行して相関性指数を更新するサブ処理と、(c-7) サブ処理(c-6)の終了後、サブ処理(c-0)に戻り第2の分子に含まれる複数の原子の中から未だ選択されていない他の原子を選択し、ここで選択された他の原子についてサブ処理(c-2)～(c-6)を実行し、相関性指数を加算して類似性指数を求めるサブ処理とを有する処理と、(d) 処理(c)の終了後、三次元座標内で第1の分子と第2の分子との相対位置を変更する処理とを実行させることを特徴とする分子類似性評価プログラム。

【請求項16】 サブ処理(c-4)で求める原子間距離は、直交座標系におけるいずれかの軸に関する距離であることを特徴とする請求項15に記載の分子類似性評価プログラム。

30 【請求項17】 処理(d)は、類似性指数が高まる方向に向けて、第1の分子と第2の分子との相対位置を変更することを特徴とする請求項15に記載の分子類似性評価プログラム。

【請求項18】 原子に固有の値は、該原子のファンデルワールス半径、原子半径、共有半径のいずれかであることを特徴とする請求項13から17のいずれか一に記載の分子類似性評価プログラム。

【請求項19】 上限閾値は、原子に固有の値に上限係数を乗じて得られる値であることを特徴とする請求項13から18のいずれか一に記載の分子類似性評価プログラム。

40 【請求項20】 下限閾値は、原子に固有の値に下限係数を乗じて得られる値であることを特徴とする請求項13から18のいずれか一に記載の分子類似性評価プログラム。

【請求項21】 上限係数が1.0～1.3の範囲内の値であることを特徴とする請求項19に記載の分子類似性評価プログラム。

50 【請求項22】 下限係数が0.5～0.8の範囲内の値であることを特徴とする請求項20に記載の分子類似

$$S_{AB} = \frac{C}{(T_A \cdot T_B)^{\frac{1}{2}}}$$

性評価プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、2つの分子の形状類似性を評価する方法とプログラムに関する。

【0002】

【従来の技術】目的とする生理活性を有する薬物分子を設計する方法として、三次元構造活性相關（3D-QSAR）解析やファーマコフォアマッピングを用いた、論理的な分子設計手法が用いられている。これらの方法は、既存薬物同士を適当なルールに従って三次元仮想空間内で重ね合わせた後、PLS (partial least square of latent valuables) 法、ニューラルネット (NN) 法、遺伝アルゴリズム (GA) などを用いて統計処理することにより、それら薬物の活性と種々のパラメータ間の特徴を抽出するものである。そして、得られた結果をグラフィックス表示することにより、分子構造中の活性に寄与する部分（例えば、官能基、立体構造）を視覚的に認識し、分子設計の手がかりとすることが可能である。また、新規にデザインした分子の活性を予測することにも応用される。

【0003】3D-QSARやファーマコフォア探索では、分子を重ね合わせるための手法が重要である。これまで、比較する複数の分子間で、対応すると予想される原子-原子あるいは官能基-官能基を最小二乗平方和平方根 (root mean square) が最小となるように漸次分子を移動して重ね合わせの度合いを高めていく方法や、評価関数（分子類似性）を用いて最も良い重なりを順次探していく方法が用いられてきた。

【0004】ところが、原子-原子又は官能基-官能基を重ね合わせる方法では、重ね合わせを短時間で行えるが、研究者の主観が入り込むことが避けられない。また、計算機により官能基を自動抽出する方法では、重ね合わせる官能基の種類又は数の選択方法に、ソフトウェア依存の任意性や研究者の主観が含まれるといった問題が残る。一方、評価関数を用いる方法では、分子の重ね合わせ手法そのものとしては理想的であるが、計算に時間がかかるという別の問題がある。

【0005】例えば、Carboらにより提案された分子類似性評価法は、分子の電子密度を比較するものであるが、多くの計算時間を必要とするものであった。そのため、新たに、分子静電ポテンシャル (MEP) や形状を比較する分子類似性の評価方法が別に提案されている。

【0006】Journal of Computer-Aided Molecular Design, 5(1991), 427-439, "Similarity of molecular shape" に開示されているMeyerらにより提案された分子形状類似性評価法では、以下の式を用いて分子類似性が評価される。

式中、 T_A 、 T_B は、それぞれ2つの分子A、Bの体積を表し、Cは分子A、Bを重ね合わせたときに両者が共有する部分の面積である。分子A、Bが完全に重なり合うとき、 S_{AB} は最大値 (+1) となり、重なり合う部分が無いとき S_{AB} には 0 が与えられる。各体積の値は、一定間隔で発生させた格子点のうち、それぞれの分子又は両者のファンデルワールス半径内に含まれる格子点を数えることにより計算される。

【0007】しかし、格子点を用いた分子類似性の評価は、計算に多くの時間を要し、正確な結果を得るために非常に細かい格子点（通常は 0.2 Å の格子点）が必要とされる。したがって、最適な分子の重ね合わせモデルを得るには、ワークステーション等の計算速度の速い計算機が必要である。

【0008】この問題を解消するために、Richardsらは、鋳型分子中の原子と比較する分子中の原子間の距離を用いたガウス関数を計算し、分子類似性の計算に応用している。Perspectives in Drug Discovery and Design, 9/10/11:321-338, 1998, "Explicit Calculation of 3D Molecular Similarity" 参照。この方法は、Meyerらの方法に比べると相当高速に分子類似性を評価できる。しかし、分子類似性の計算には、一方の分子のすべての原子について他方のすべての原子との間の原子間距離を計算する必要があり、依然として計算時間のかかるものである。

【0009】以上のように、従来の分子類似性評価法はいずれも多くの計算時間がかかることから、繰返し分子類似性を求め、その結果をもとにシンプレスクス法などで最適な重ね合わせモデルを簡便に得ることができなかった。また、一つの官能基と別の官能基との距離が一定距離内になるように重ね合わせる方法は、ファーマコフォアが既に明らかな場合には特に威力を発揮するが、官能基の数が少ない低分子では、良好な重ね合わせモデルを得ることができない。

【0010】

【発明が解決しようとする課題】そこで、本発明は、簡便で且つ高速な分子類似性評価法を提供すること、また類似性が最大となるように分子を移動させて最適な重ね合わせモデルを得る方法、さらにそれらの方法を実行するプログラムを提供するものである。

【0011】

【課題を解決するための手段】このような課題を解決するために、本発明に係る分子類似性評価方法は、第1の分子と第2の分子との類似性を評価するものであって、

(a) 第1の分子に含まれる原子に固有の値、又は該原子との相関性を評価する第2の分子の原子に固有の値、若しくはそれらの値から求まる別の値のうちのいず

れかの値をもとに上限閾値と下限閾値とを求めるステップと、(b) 第1の分子の原子と第2の分子の原子との相関性を上限閾値と下限閾値とを用いて計算するステップと、(c) ステップ(b)で得られた相関性をもとに第1の分子と第2の分子の類似性を評価するステップとを備えたものである。

【0012】本発明の他の形態の分子類似性評価方法は、第1の分子に含まれる各原子と第2の分子に含まれる各原子との原子間距離をもとに原子間の相関性を求め、この相関性から第1と第2の分子の類似性を評価するものであって、2つの原子間の相関性を求める処理に、(a) 第1の分子の原子及び／又は第2の分子の原子に固有の値を求めるステップと、(b) ステップ(a)で求めた固有の値から上限閾値を求めるステップと、(c) ステップ(a)で求めた固有の値から下限閾値を求めるステップと、(d) 2つの原子間の距離を求めるステップと、(e) ステップ(d)で求めた原子間の距離と、上限閾値と、下限閾値とを用いて、第1の分子の原子と第2の分子の原子との相関性を評価するステップとを含むものである。

【0013】本発明の他の形態は、三次元座標系に配置された第1の分子と第2の分子との類似性を評価する方法及びプログラムに、(a) 相関性指数をリセットするステップと、(b) 第2の分子に含まれる複数の原子の中から一つの原子を選択するステップと、(c) 第1の分子に含まれる複数の原子の中から一つの原子を選択するステップと、(d) ステップ(b)と(c)でそれぞれ選択された原子の双方又は一方に固有の値から上限閾値と下限閾値とを求めるステップと、(e) ステップ(b)と(c)でそれぞれ選択された原子間距離を求めるステップと、(f) ステップ(d)で求めた上限閾値と下限閾値をもとに、ステップ(e)で求めた原子間距離に応じた相関性指数を求めるステップと、(g) ステップ(f)の終了後、第1の分子に含まれる複数の原子の中から未だ選択されていない原子が有るか否か判断するステップと、(h) ステップ(g)で、第1の分子に含まれる複数の原子の中から未だ選択されていない原子が有ると判断されると、ステップ(c)に戻り第1の分子に含まれる複数の原子の中から未だ選択されていない他の原子を選択し、ステップ(d)～(f)を実行して相関性指数を更新するステップと、(i) ステップ(g)で、第1の分子に含まれる複数の原子の中から未だ選択されていない原子が無いと判断されると、ステップ(a)に戻り第2の分子に含まれる複数の原子の中から未だ選択されていない他の原子を選択し、ステップ(c)～(h)を実行し、相関性指数を加算するステップとを有するものである。なお、ステップ(d)は、類似性指数が高まる方向に向けて、第1の分子と第2の分子との相対位置を変更するものであることが好ましい。

【0014】本発明の他の形態は、2つの分子の類似性を評価する方法及びプログラムに、(a) 第1及び第2の分子に含まれる各原子の三次元座標を得るステップと、(b) 第1及び第2の分子を三次元座標系に配置するステップと、(c) 第2の分子に含まれる各原子について第1の分子に含まれる各原子との類似性指数を求めるステップであって、(c-0) 相関性指数をリセットするサブステップと、(c-1) 第2の分子に含まれる複数の原子の中から一つの原子を選択するサブステップと、(c-2) 第1の分子に含まれる複数の原子の中から一つの原子を選択するサブステップと、(c-3) サブステップ(c-1)と(c-2)でそれぞれ選択された原子の双方又は一方に固有の値から上限閾値と下限閾値とを求めるサブステップと、(c-4) サブステップ(c-1)と(c-2)でそれぞれ選択された原子間距離を求めるサブステップと、(c-5) サブステップ(c-3)で求めた上限閾値と下限閾値をもとに、サブステップ(c-4)で求めた原子間距離に応じた相関性指数を求めるサブステップと、(c-6) サブステップ(c-5)の終了後、第1の分子に含まれる複数の原子の中から未だ選択されていない原子が有るか否か判断するステップと、(c-7) サブステップ(c-6)で、第1の分子に含まれる複数の原子の中から未だ選択されていない原子が有ると判断されると、サブステップ(c-2)に戻り第1の分子に含まれる複数の原子の中から未だ選択されていない他の原子を選択し、サブステップ(c-3)～(c-6)を実行して相関性指数を更新するステップと、(c-8) サブステップ(c-6)で、第1の分子に含まれる複数の原子の中から未だ選択されていない原子が無いと判断されると、ステップ(c-0)に戻り第2の分子に含まれる複数の原子の中から未だ選択されていない他の原子を選択し、ステップ(c-2)～(c-7)を実行し、相関性指数を加算して類似性指数を求めるサブステップとを有するステップと、(d) ステップ(c)の終了後、三次元座標内で第1の分子と第2の分子との相対位置を変更するステップとを含むものである。

【0015】以上の複数の形態において求める原子間距離は、直交座標系におけるいざれかの軸に関する距離であつてもよい。

【0016】また、原子に固有の値は、該原子のファンデルワールス半径、原子半径、共有半径のいざれかを利用できる。

【0017】さらに、上限閾値は原子に固有の値に上限係数を乗じて得られる値、また下限閾値は原子に固有の値に下限係数を乗じて得られる値であることが好ましい。そして、上限係数は1.0～1.3の範囲内の値、下限係数は0.5～0.8の範囲内の値が好適に利用できる。

50 【0018】

【発明の実施の形態】添付図面を参照して、重ね合わせ法を用いた2つの分子の類似性評価について説明する。なお、以下に説明するプロセスは、本発明に基づいてコーディングされたプログラムを適当なコンピュータ上で起動させることにより行われるものである。当然、コンピュータは、プログラムの実行に必要な種々の装置（例えば、CPU、内部・外部メモリ等）を含むものである。プログラムは、現在提供されている種々のオペレーティングシステム上で動作するものであり、システムの構成や該システムに搭載されているアプリケーション等は本発明にとって制限的なものでない。また、プログラムは、CD-ROM等の種々の既存の記録媒体に記録され、又はインターネットや電話回線等の通信回線を通じて、提供される。

【0019】類似性評価について、図1～図4に示す第*

表-1

分子	原子	X座標	Y座標	Z座標
M(1)	A(1,1)	x11	y11	z11
	A(1,2)	x12	y12	z12
	A(1,3)	x13	y13	z13
M(2)	A(2,1)	x21	y21	z21
	A(2,2)	x22	y22	z22
	A(2,3)	x23	y23	z23
	A(2,4)	x24	y24	z24

【0020】ステップ3では、第2の分子M(2)の選択された一つの原子について、第1の分子M(1)の選択された一つの原子との距離を計算する。なお、便宜上、原子数の多い第2の分子を鑄型分子、原子数の少ない第1の分子を比較分子、鑄型分子に含まれる原子を鑄型原子、比較分子に含まれる原子を比較原子と必要に応じていう。本実施形態の場合、第2の分子M(2)が鑄型分子、第1の分子M(1)が比較分子となる。

【0021】例えば、図示する例では、選択鑄型原子A(2,1)と比較原子A(1,1)との距離r(21,11)を計算する。

【0022】原子間距離は、以下の一般式(1)で与えられる。

$$r = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (1)$$

x1, y1, z1: 比較分子の座標
x2, y2, z2: 鑄型分子の座標

【0023】ステップ4では、ステップ3で計算された原子間距離r(21,11)が、鑄型原子A(2,1)と比較原子A(1,1)について固有の第1の閾値(上限閾値)rlim1(1,1)以下か否か判断する。第1の閾値rlim1(1,1)は、例えば、下記の式(2)に示すように、選択鑄型原子A(2,1)と比較原子A

*1の分子M(1)と第2の分子M(2)を例にとり、図5のフローチャートを参照して説明する。まず、本発明の類似性評価法では、各分子M(1)、M(2)に含まれる原子の三次元座標値(例えば、X、Y、Z座標)を入手する(ステップ1)。次に、入手した三次元座標値を利用し、三次元座標仮想空間(直交座標空間)に2つの分子M(1)、M(2)を配置する(ステップ2)。第1の分子M(1)が3原子[A(1,1)、A(1,2)、A(1,3)]から構成され、第2の分子M(2)が4原子[A(2,1)、A(2,2)、A(2,3)、A(2,4)]から構成されており、三次元座標空間に配置された2つの分子M(1)、M(2)の各原子が、以下の表-1に示す座標を与えられているものとする。

10

(1,1)のファンデルワールス半径vdw[A(2,1)]、vdw[A(1,1)]の平均値に所定の第1の係数k1を乗じた値が好ましく用いられる。なお、第1の閾値rlim1(1,1)は、上述したステップ3で計算することができる。

$$rlim(1,1) = k1 \cdot \frac{vdw[A(2,1)] + vdw[A(1,1)]}{2} \quad (2)$$

【0024】第1の閾値は、2つの原子のいずれか一方の原子のファンデルワールス半径、又は2つの原子のファンデルワールス半径のうち大きい値、若しくは2つの原子のファンデルワールス半径のうち小さい値、若しくはそれらの値に第1の係数を乗じた値であってもよい。

【0025】なお、鑄型原子のX、Y又はZ座標値と比較原子のX、Y又はZ座標値との差(例えば、x21-x1、y21-y1又はz21-z1)が第1の閾値rlim1(1,1)よりも大きければ、三次元空間内における2つの原子間距離は当然第1の閾値rlim1(1,1)以上である。したがって、原子間距離を計算する前に、鑄型原子のX、Y、Z座標値と比較原子のX、Y、Z座標値との差(x21-x1、y21-y1、z21-z1)を計算し、これら3つの差がいずれも第1の閾値rlim1(1,1)以下の場合のみ、原子間距離を計算してもよい。この場合、無駄な計算が少くなり、計算速度が一段と速くなり、また50 必要なメモリが少なくなる。

【0026】原子間距離が第1の閾値 r_{lim1} (1, 1)を越える場合、以下に説明するステップ5, 6を越えてステップ7に進む。原子間距離が第1の閾値 r_{lim1} (1, 1)以下の場合、次のステップ5に進む。

【0027】ステップ5では、ステップ3で計算された原子間距離 r (21, 11)が、鋳型原子A (2, 1)と比較原子A (1, 1)について固有の第2の閾値(下限閾値) r_{lim2} (1, 1)以下か否か判断する。第2の閾値 r_{lim2} (1, 1)は、例えば、下記の式(3)に示すように、選択鋳型原子A (2, 1)と比較原子A (1, 1)のそれぞれのファンデルワールス半径の平均値に所定の第2の係数 k_2 を乗じた値を用いる。この第2の閾値 r_{lim2} (1, 1)は、上述したステップ3で計算することができる。なお、第2の係数 k_2 は、第1の閾値 r_{lim1} (1, 1)を求める際に用いた第1の係数 k

表-2

分岐条件	分岐先	分岐先処理
$r_{lim1} < r$	S 7	si, sitemp : 変化なし
$r_{lim2} \leq r < r_{lim1}$	S 11	sitemp ← 0. 5
$r \leq r_{lim2}$	S 6	si ← si + 1

表中、「si」は、鋳型分子と比較分子との重なり度合い、すなわち両分子の類似性を示す指標であり、この類似性指数が高ければ高いほど両者の類似性は高いことを意味する。「sitemp」は、選択された鋳型原子と比較分子との相関性を示す指標であり、すべての鋳型原子について個々に得られた相関性指数sitempの総和が、鋳型分子と比較分子との類似性siに相当する。なお、類似性指数si、相関性指数sitempは、ステップ3に進む前のいずれかのステップ(例えば、ステップ2)で、初期値「0」にリセットされる。

【0031】分岐処理について具体的に説明する。例えば、原子間距離 r (21, 11)が第1の閾値(上限閾値) r_{lim1} (1, 1)を越えると判断されたとする。この場合、鋳型原子A (2, 1)と比較原子A (1, 1)は三次元空間上で離れており、鋳型原子A (2, 1)と比較分子M (1)との相関性は低い。したがって、類似性指数si、相関性指標sitempは変化しない。

【0032】原子間距離 r (21, 11)が第1の閾値 r_{lim1} (1, 1)以下で第2の閾値 r_{lim2} (1, 1)を越える場合、鋳型原子A (2, 1)と比較原子A (1, 1)は三次元空間上で比較的近い位置にあり、鋳型分子M (2)と比較分子M1 (1)との相関性、類似性を高める要因となる。したがって、ステップ6で相関性指数sitempに所定のロースコア(例えば、「0. 5」)を与える。

【0033】原子間距離 r (21, 11)が第2の閾値 r_{lim2} (1, 1)以下の場合、鋳型原子A (2, 1)と比較原子A (1, 1)は三次元空間上で非常に接近しており、鋳型分子M (2)と比較分子M1 (1)との相関

1よりも小さい値である。

$$r_{lim2}(L1) = k_2 \cdot \frac{vdw[Atom(2,1)] + vdw[Atom(L1)]}{2} \quad (3)$$

【0028】第2の閾値は、2つの原子のいずれか一方の原子のファンデルワールス半径、又は2つの原子のファンデルワールス半径のうち大きい値、若しくは2つの原子のファンデルワールス半径のうち小さい値、若しくはそれらの値に所定の第2の係数を乗じた値であつてもよい。

10 【0029】周知の通りファンデルワールス半径は各原子に固有の値である。したがって、第1及び第2の閾値は、計算対象となる2つの原子の種類に応じて異なる。

【0030】ステップ4, 5での判断結果に応じて、プログラムの処理は次の表に示すルートに分岐する。

性、類似性を高める大きな要因となる。したがって、ステップ11で類似性指数siに所定のハイスコア(例えば、「1」)を加算する。

【0034】鋳型原子A (2, 1)と比較原子A (1, 1)は三次元空間上で非常に接近している場合(すなわち、類似性指数siにハイスコア「1」が加算された場合)、選択鋳型原子A (2, 1)とその他の比較原子A (1, 2)、A (1, 3)との原子間距離は、鋳型原子A (2, 1)と比較原子A (1, 1)との原子間距離以上と考えるのが妥当である。したがって、この場合、鋳型原子選択鋳型原子A (2, 1)とその他の比較原子A (1, 2)、A (1, 3)との原子間距離は計算せず、選択された鋳型原子と比較分子との相関性指数sitempに変更を加えず、その類似性指数siにハイスコア「1」を加算し、当該鋳型原子に関する原子間距離の計算は終了する。

【0035】一方、鋳型原子A (2, 1)と比較原子A (1, 1)が比較的接近している場合でも、他の鋳型原子A (2, 2)、A (2, 3)又はA (2, 4)が鋳型原子(2, 1)よりも更に比較原子A (1, 1)に接近している可能性がある。したがって、この場合、類似性指数siは変更せず、取り敢えず相関性指標sitempにロースコア「0. 5」を与える。

【0036】ただし、ハイスコア、ロースコアの値はそれ「1」、「0. 5」に限るものではなく、本プログラムの利用者が適宜設定できるようにしてよい。

40 【0037】ステップ11で類似性指数siにハイスコア「1」が加算された場合、上述のように、2つの原子A (2, 1)とA (1, 1)が空間座標内で極近傍に位置

しており、もはや他の比較原子A(1, 2)、A(1, 3)との原子間距離は計算する必要がないので、鋳型分子M(2)の他の鋳型原子A(2, 2)と比較分子M(1)の原子A(1, 1)との原子間距離の計算に移る(ステップ10)。

【0038】ステップ6で相関性指数sitempにロースコア「0.5」が付与された場合、ステップ7に進む。ステップ7では、選択鋳型原子A(2, 1)と他の比較原子A(1, 2)、(1, 3)について、上述したステップ3～6の処理が終了しているか否か判断する。比較分子のすべての原子について上記手順が終了していないときはステップ3に戻り第1の分子M(1)の中から次の比較原子A(1, 2)を選択し、この選択された比較原子A(1, 2)について上述のステップを繰り返す。その際、ステップ4, 5で原子間距離r(21, 12)が第1の閾値rlim2(1, 2)及び第2の閾値rlim2(1, 2)以下と判断されると、類似性指数siにハイスコア「1」を加算する。また、残りの比較原子A(1, 3)との原子間距離は計算せず、鋳型分子の他の鋳型原子A(2, 2)と比較分子の各原子A(1, 1)との原子間距離の計算に移る。

【0039】A(1, 3)についても、比較原子A(1, 1)、A(1, 2)と同様に、鋳型原子A(2, 1)との原子間距離がr(21, 13)が第1の閾値rlim1(1, 3)以下で第2の閾値rlim2(1, 3)を越えると判断された場合、ステップ6で相関性指数sitempに所定のロースコア「0.5」を与える。つまり、相関性指数sitempは「0.5」の状態に維持される。ただし、原子間距離r(21, 13)が第1の閾値rlim2(1, 3)及び第2の閾値rlim2(1, 3)以下と判断されると、類似性指数siにハイスコア「1」を加算する。

【0040】以上で説明したように、選択された鋳型原子について各比較原子との間でステップ4～7の処理を繰り返している間、原子間距離が第1の閾値以下で且つ第2の閾値を越えると判断されると、その選択された鋳型原子に対する相関性指数sitempはロースコア「0.5」に維持される。また、選択された鋳型原子について一旦相関性指数sitempとしてロースコア「0.5」が付与されると、当該選択された鋳型原子に関する計算の過程で原子間距離が第1の閾値以下で且つ第2の閾値を越えると判断されても、また原子間距離が第1の閾値を越えると判断されても、そのロースコアが維持される。ただし、鋳型原子について、ステップ5で原子間距離が第1及び第2の閾値以下と判断された場合、類似性指数siにハイスコア「1」が加算される。

【0041】ステップ7で、選択された鋳型原子について、すべての比較原子との原子間距離の計算が終了したと判断されると、ステップ8で、類似性指数siに相関性指数sitemp(ロースコア「0.5」)を加算する。次に、ステップ9で、相関性指数sitempを「0」にリセット

として次の選択鋳型原子に対する計算に備える。

【0042】ステップ10では、鋳型分子M(2)のすべての鋳型原子A(2, 1)～A(2, 4)について上述の処理が終了したか否か判断し、未終了であればステップ12で、鋳型分子の他の原子を選択してステップ3に戻る(図3, 4参照)。

【0043】以上のステップ3～11の処理により、2つの分子の一つの三次元空間配置について、これら2つの分子の類似性指数siが計算される。

- 10 【0044】次に、ステップ12では、一つの三次元空間配置に対して得られた類似性指数siが最大値か否か判断する。最大値でない判断されると、三次元空間内における第1と第2の分子の相対的位置関係(例えば、回転、平行移動)に変更を加えた(例えば、回転、平行移動した)後、上述した処理を繰り返す。第1又は第2の分子の移動は、X, Y, Zの3軸に平行な方向への移動のみならず、これらX, Y, Z又はそれらと平行な軸を中心とする回転が含まれる。また、類似性指数siが最大化されるように分子を移動することは非常に重要な技術であり、それについてはNelder, J.A., and Mead, R. 1965, Computer Journal, vol. 7, p. 308に記載されたシンプレックス法(Simplex method)があり、そのソースコードの一例がNumerical Recipes in C: The Art of Scientific Computing, The Press Syndicate of the University of Cambridge(日本語版:ニューメリカル・レシピ・イン・C, 株式会社技術評論社)に記載されている。

- 20 【0045】なお、以上の説明では、第1及び第2の閾値は、原子のファンデルワールス半径をもとに決定したが、ファンデルワールス半径以外に、原子半径、共有半径、又はそれらを組み合わせた値が利用できる。

- 30 【0046】また、原子の座標値は直交座標系で表す必要はなく、極座標で表してもよい。この場合、三次元空間は極座標で与えられる。

- 【0047】さらに、上述のようにして最大の類似性指数が得られた分子の配置に対して、Meyerら又はRichards等によって提案されている分子類似性評価法を適用すれば、更に高精度に分子の類似性を評価できる。Journal of Computer-Aided Molecular Design, 5(1991), 427-439, "Similarity of molecular shape", Perspectives in Drug Discovery and Design, 9/10/1:321-338, 1998, "Explicit Calculation of 3D Molecular Similarity"参照。

【0048】

【実施例1】類似性評価の具体的な実施例を説明する。

- 【0049】1. 評価対象: 本発明の理解を容易にするため、図6と図7に示す分子構造の簡単な水分子M(1)とアンモニア分子M(2)についてそれらの類似性を評価する。ここで、水分子M(1)の原子数は「3」、アンモニア分子M(2)の原子数は「4」であ

る。したがって、アンモニア分子M(2)が鋳型分子、水分子M(1)が比較分子となる。

【0050】2. 原子座標

水分子とアンモニア分子のそれについて、分子に含まれる原子の三次元データ（三次元座標、分子間距離、結合角度等）はChem3D(CambridgeSoft社)などで作成する。Chem3Dなどの分子エディタで作成した後、MM2やAMBERなどの分子力場計算やMOPACなどの分子軌道計算を行い、分子の最安定構造を求めることがある。また、PDB(プロテインデータバンク)などに登録されている分子の三次元構造を利用することもある。タキソールはMSI*

*社のWebLab ViewerLiteの座標を、タクロリムスはPDBからダウンロードした座標を用いている。また、水とアンモニアは、Chem3Dで構造を作成した。

【0051】3. 分子の配置

入手した水分子とアンモニア分子の三次元座標をもとに、図6に示すように、三次元座標系の中に、水分子とアンモニア分子を配置する。

【0052】原子の配置座標を以下の表3に示す。なお、表-3及び以下の説明中に用いる座標値の単位はすべてÅである。

表-3 [単位: Å]

分子	原子	X座標	Y座標	Z座標
水: M(1)	O(1,1)	-0.151	+0.236	-0.075
	H(1,2)	-0.156	+1.166	+0.075
	H(1,3)	+0.759	+0.003	-0.011
アンモニア:	N(2,1)	-0.121	-0.144	-0.398
	H(2,2)	+0.687	-0.766	-0.380
M(2)	H(2,3)	+0.249	+0.766	-0.124
	H(2,4)	-0.687	-0.439	+0.398

【0053】4. 類似性評価

鋳型分子であるアンモニア分子の各原子について、水分子における各原子との原子類似性を求める。具体的には、以下の手順(処理)に従う。

【0054】(1) 手順1

窒素原子N(2,1)と酸素原子O(1,1)との距離と、X、Y、Zのいずれかの座標に関してこれら2つの原子のファンデルワールス半径から定まる第1の閾値 $r_{lim1}(21,11)$ と比較する。

【0055】第1の閾値は、計算に係る2つの原子のそれぞれのファンデルワールス半径[vdw]の平均値に第1の係数 $k_1 = 1.3$ を乗じた値として与えられる。ただし、第1の係数は、この値に限るものでない。具体的に、窒素原子N(2,1)と酸素原子O(1,1)に関して、窒素のファンデルワールス半径[vdw(N)]

$$Abs(x_2 - x_1) = -0.121 - (-0.151) = 0.030 < 1.911 \quad (7)$$

$$Abs(y_2 - y_1) = -0.144 - (+0.236) = 0.380 < 1.911 \quad (8)$$

$$Abs(z_2 - z_1) = -0.398 - (-0.075) = 0.323 < 1.911 \quad (9)$$

【0057】式(7)～(9)から明らかなように、窒素原子N(2,1)と酸素原子O(1,1)に関してX、Y、Z座標の差は、第1の閾値 $r_{lim1}(1,1)$ 以下である。したがって、アンモニア分子M(2)の窒素原子N(2,1)と、水分子M(1)の酸素原子O(1)との距離 $r(21,11)$ を演算する。上述のように、窒素原子N(2,1)と酸素原子O(1,1)の三次元座標は既知であるから、その三次元座標を用いると次の式(10)の通り両者の原子間距離 $r(21,11)$ が求

と酸素のファンデルワールス半径[vdw(O)]はそれぞれ以下の式(4)、(5)の通り与えられるので、第1の閾値 $r_{lim1}(21,11)$ は式(6)の通り計算される。

$$vdw(N) = 1.54 \quad (4)$$

$$vdw(O) = 1.40 \quad (5)$$

$$r_{lim1}(21,11) = 1.3 \cdot \left(\frac{1.54 + 1.40}{2} \right) = 1.91 \quad (6)$$

【0056】次に、以下の式(7)～(9)に示すように、窒素原子N(2,1)と酸素原子O(1,1)に関してX、Y、Z座標の差を求める。なお、式(7)～(9)には、第1の閾値 $r_{lim1}(1,1)$ と比較結果も併せて表示してある。

まる。

$$r(21,11) = \sqrt{(0.030)^2 + (0.380)^2 + (0.323)^2} = 0.5 \quad (10)$$

【0058】計算された原子間距離 $r(21,11)$ と、第2の閾値との大小を比較する。第2の閾値 $r_{lim2}(21,11)$ は、計算に係る2つの原子のそれぞれのファンデルワールス半径[vdw]の平均値に第2の係数 $k_2 = 0.6$ を乗じた値として与えられる。なお、第2の係数はこの値に限るものでない。具体的に、第2の

閾値 $r_{lim1}(21, 11)$ は次の式(11)の通り計算される。

$$r_{lim2}(21,11)=0.6 \cdot \left(\frac{1.54+1.40}{2} \right) = 0.882 \quad (11)$$

【0059】以上のようにして求めた第1の閾値 r_{lim1} *

$$r(21,11)=0.500 < r_{lim2}(21,11)=0.882 < r_{lim1}(21,11)=1.911 \quad (12)$$

したがって、窒素原子N(1, 1)と酸素原子O(1, 1)との類似性指数 s_i にはハイスクア「1」を与える。この場合、もはや窒素原子N(2, 1)と水分子M(1)の他の水素原子H(1, 2)、H(1, 3)との原子間距離は計算せず、アンモニア分子M(2)の水素原子H(2, 2)と水分子の各原子との計算に移る。

【0060】(2)手順2

① アンモニア分子M(2)の水素原子H(2, 2)について、以下の式(13)～(15)の通り水分子M(1)の酸素原子O(1, 1)とのX、Y、Z座標差を求める。

$$Abs(x_2 - x_1) = +0.687 - (-0.151) = 0.838 \quad (13)$$

$$Abs(y_2 - y_1) = -0.766 - (+0.236) = 1.002 \quad (14)$$

$$Abs(z_2 - z_1) = -0.380 - (-0.075) = -0.305 \quad (15)$$

【0061】計算で得られた座標差を第1の閾値 $r_{lim1}(22, 11)$ 、第2の閾値 $r_{lim2}(22, 11)$ と比較する。なお、第1の閾値 $r_{lim1}(22, 11)$ 、第2の閾値 $r_{lim2}(22, 11)$ はそれぞれ以下の式

$$r_{lim2}(22,11)=0.78 < r(22,11)=1.340 < r_{lim1}(22,11)=1.690 \quad (19)$$

【0064】したがって、水原子H(2, 2)について相関性指數 $sitemp(2)$ として「0. 5」を与える。

【0065】② アンモニア分子M(2)の水素原子H(2, 2)について、水分子M(1)の水素原子H(1, 2)との処理を、手順2と同様に行う。なお、以下では、記載を簡略化するために、各計算の説明は最小限に止める。まず、アンモニア分子M(2)の水素原子H(2, 2)について、水分子M(1)の水素原子H(1, 2)とのX、Y、Z座標差を以下の式(20)～(22)の通り求める。

$$Abs(x_2 - x_1) = +0.687 - (+0.155) = 0.842 \quad (20)$$

$$Abs(y_2 - y_1) = -0.766 - (+1.116) = 1.822 \quad (21)$$

$$Abs(z_2 - z_1) = -0.380 - (+0.075) = -0.455 \quad (22)$$

第1の閾値 $r_{lim1}(22, 12)$ 、第2の閾値 $r_{lim2}(22, 12)$ は、次の式(23)、(24)で計算される通りである。

$$r_{lim1}(22,12)=1.3 \cdot \left(\frac{1.20+1.20}{2} \right) = 1.560 \quad (23)$$

$$r_{lim2}(22,12)=0.6 \cdot \left(\frac{1.20+1.20}{2} \right) = 0.720 \quad (24)$$

計算の結果を見れば明らかなように、Y座標値差(1.822)は、第1の閾値 $r_{lim1}(22, 12) = 1.$

* (21, 11)、第2の閾値 $r_{lim2}(21, 11)$ と、窒素原子N(2, 1)と酸素原子O(1, 1)の原子間距離 $r(21, 11)$ は、以下の式(12)の関係を有する。

(16)、(17)の通り計算される。

$$r_{lim1}(21,11)=1.3 \cdot \left(\frac{1.20+1.40}{2} \right) = 1.690 \quad (16)$$

$$r_{lim2}(21,11)=0.6 \cdot \left(\frac{1.20+1.40}{2} \right) = 0.780 \quad (17)$$

【0062】計算の結果を見れば明らかなように、いずれの座標差(0.838、1.002、-0.305)も第1の閾値 $r_{lim1}(22, 11) = 1.690$ 以下である。したがって、以下の式(18)の通り、水素原子H(2, 2)と酸素原子O(1, 1)との原子間距離 $r(22, 11)$ を計算する。

$$r(22,11)=\sqrt{(0.838)^2+(1.002)^2+(0.305)^2}=1.340 \quad (18)$$

【0063】計算で得られた原子間距離 $r(22, 11)$ は、第1の閾値 $r_{lim1}(21, 11)$ 、第2の閾値 $r_{lim2}(21, 11)$ と以下の式(19)に示す関係にある。

$$r_{lim2}(22,11)=0.78 < r(22,11)=1.340 < r_{lim1}(22,11)=1.690 \quad (19)$$

560よりも大きい。したがって、相関性指數 $sitemp(2)$ を「0. 5」に維持し、次の処理に進む。

【0066】③ アンモニア分子M(2)の水素原子H(2, 2)について、水分子M(1)の水素原子H(1, 3)との処理を同様に行う。まず、水素原子H(2, 2)について、水分子M(1)の水素原子H(1, 3)とのX、Y、Z座標差を以下の式(25)～(27)の通り求める。

$$Abs(x_2 - x_1) = +0.687 - (+0.759) = -0.072 \quad (25)$$

$$Abs(y_2 - y_1) = -0.766 - (+0.003) = -0.769 \quad (26)$$

$$Abs(z_2 - z_1) = -0.380 - (-0.011) = -0.369 \quad (27)$$

第1の閾値 $r_{lim1}(22, 13)$ 、第2の閾値 $r_{lim2}(22, 13)$ は、次の式(28)、(29)で計算される。

$$r_{lim1}(22,13)=1.3 \cdot \left(\frac{1.20+1.20}{2} \right) = 1.560 \quad (28)$$

$$r_{lim2}(22,13)=0.6 \cdot \left(\frac{1.20+1.20}{2} \right) = 0.720 \quad (29)$$

X、Y、Z座標差はいずれも第1の閾値 $r_{lim1}(22, 13)$ 以下であるので、次の式(30)より原子間距離 $r(22, 13)$ を求める。

$$r(22,13) = \sqrt{(0.072)^2 + (0.769)^2 + (0.369)^2} = 0.856 \quad (30)$$

原子間距離 $r(22, 13)$ は、第1の閾値 $r_{lim1}(2*$

$$r_{lim2}(22,13) = 0.72 < r(22,13) = 0.856 < r_{lim1}(22,13) = 1.560 \quad (31)$$

したがって、水原子 H(2, 2)について相関性指数 $sitemp(2) = '0.5'$ を維持する。

【0067】(3) 手順3

同様にしてアンモニア分子 M(2) の水素原子 H(2, 3) について、水分子 M(2) の各原子との相関性指数 $sitemp(3)$ を求める。計算の過程は省略するが、相関性指数 $sitemp(3) = 1$ と計算される。

【0068】(4) 手順4

$$si = sitemp(1) + sitemp(2) + sitemp(3) + sitemp(4) = 1 + 0.5 + 1 + 0.5 = 3 \quad (32)$$

【0070】(6) 手順6

三次元空間内で、アンモニア分子に対する水分子の位置を所定量だけ移動（座標軸と平行に移動又は回転）し、以上の手順 1～5 に基づいて類似性指数 SI を計算する。分子の移動は、まず例えばアンモニア分子に対して水分子を複数の方向に移動し、それについて相関性指数を計算する。次に、計算結果をもとに、より高い相関性指数の得られた方向に水分子を移動する。このような処理を繰り返し、類似性指数 SI の最大値が得られた段階で計算を終了する。そして、以下の式 (33) に示すように、最大の類似性指数 SI を鋳型分子の原子数 N_{mol} で除し、水分子とアンモニア分子の類似性 SI を得る。

表-4：最適係数

k1	1.3	1.1	1.2	1.3	1.3	1.0	1.1	1.2	1.0	M法	R法
k2	0.5	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.8		
SI	.974	.984	.984	.987	.986	.979	.982	.982	.981	(1.)	.993

この表4に示すように、第1の係数 k_1 を 1.3、第2の係数 k_2 を 0.6 に設定した場合に、最も高い類似性 SI が得られることが確認された。ただし、第1の係数 k_1 を 1.0～1.3、第2の係数 k_2 を 0.5～0.8 の範囲で任意に変更しても、0.97 以上の類似性が得られる。

★

★【0072】(2) 係数の対比

最適化した係数 ($k_1 = 1.3$, $k_2 = 0.6$) を用い、ランダムに配座 (400配座標) を発生させたアルドステロンと鋳型となるアルドステロンとの分子類似性の計算結果を比較した。計算結果を、表5及び図8に示す。

表-5：類似性及び標準偏差

	本発明	M法	R法
類似性: SI	0.971	(1.000)	0.977
標準偏差: SD	0.043	(0.000)	0.038

この表から明らかなように、本発明の方法と Richards らの方法の計算結果を比較すると明らかなように、本発明の方法によれば Richards らの方法に近い計算結果が得られる。

【0073】(3) タキソールの重ね合わせ

Simplex 法を用い、本発明の方法 (A)、Meyer らの方法 (B)、Richards らの方法 (C)、及び本発明の方法と Richards らの方法とを組み合わせた方法 (D) により、タキソール 2 分子を重ね合わせた。結果を表6に示す。

表6 : Simplex法によるタキソール重ね合わせ結果の対比

	本発明	Mayer	Richards	本発明 + Richards
分子類似性	0.938 (0.919)	0.986	1.000	1.000
繰返計算回数	1,219	1,743	3,633	1,285
計算時間(秒)	2.15	18,783.86	1,003.16	20.62
計算時間比	1.15×10^{-4}	1.00	5.35×10^{-1}	1.10×10^{-2}
計算時間／回(秒)	0.002	10.777	0.284	-

括弧内の値は、本発明の方法で分子の重ねあわせを最適化した後、最適化構造の分子類似性を Richards らの方法で再計算したものである。表6から明らかのように、本発明の方法によれば、計算時間を Meyer らの方法の約 5,000 倍、Richards らの方法の

約 100 倍まで高速化できる。

【0074】(4) タクロリムスの重ね合わせ
同様に、タクロリムスを重ね合わせた。結果を、表7に示す。

表7 : Simplex法によるタクロリムスの重ね合わせ結果の対比

	本発明	Mayer	Richards	本発明 + Richards
分子類似性	1.000 (0.961)	0.985	0.996	1.000
繰返計算回数	1,217	1,679	2,674	1,269
計算時間(秒)	3.09	18,257.80	1,109.01	29.02
計算時間比	1.69×10^{-4}	1.00	6.07×10^{-2}	1.59×10^{-3}
計算時間／回(秒)	0.003	11.563	0.415	-

タキソールの場合と同様の効果が確認された。

【0075】なお、以上の説明では、本発明の分子類似性評価法を形状類似性の評価に適用した例を示したが、本発明は静電容量、水素結合、疎水性を利用した分子類似性の評価法にも応用可能である。

【0076】

【発明の効果】以上の説明から明らかのように、本発明に係る分子類似性評価方法及びプログラムによれば、2つの分子の類似性を、Merer ら及び Richards らの方法よりもより少ない計算時間及び計算量で、それらの方法とほぼ同等程度に計算できる。

【図面の簡単な説明】

【図1】 本発明に係る分子類似性評価方法を説明するための図。

【図2】 図1と共に本発明に係る分子類似性評価方法を説明するために用いる分子モデル図。

【図3】 図1及び図2と共に本発明に係る分子類似性

30 評価方法を説明するために用いる分子モデル図。

【図4】 図1から図3と共に本発明に係る分子類似性評価方法を説明するために用いる分子モデル図。

【図5】 本発明に係る分子類似性評価方法及びそのプログラムの概略フローチャート。

【図6】 実施例1において分子類似性を評価した分子モデルを示す図。

【図7】 図6と共に実施例1において分子類似性を評価した分子モデルを示す図。

【図8】 実施例2の評価結果のグラフを示す図。

40 【符号の説明】

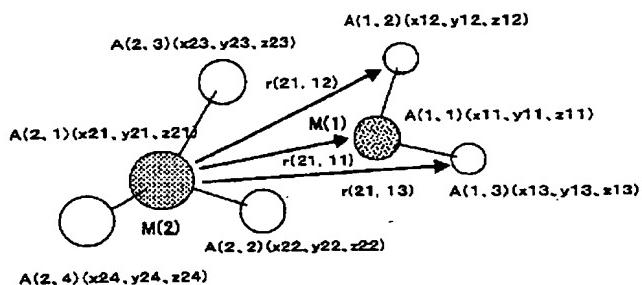
M (1) : 第1の分子

M (2) : 第2の分子

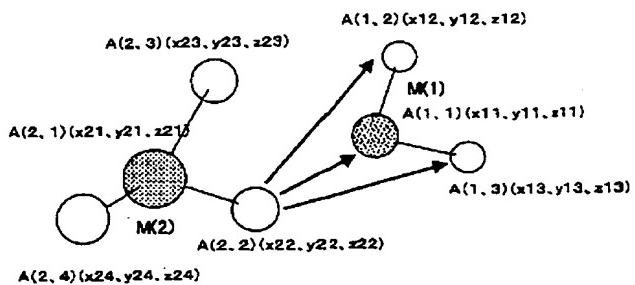
A (1, 2) ~ A (1, 3)、A (2, 1) ~ A (2, 4) : 原子

r (21, 11)、r (21, 12) r (21, 3) : 原子間距離

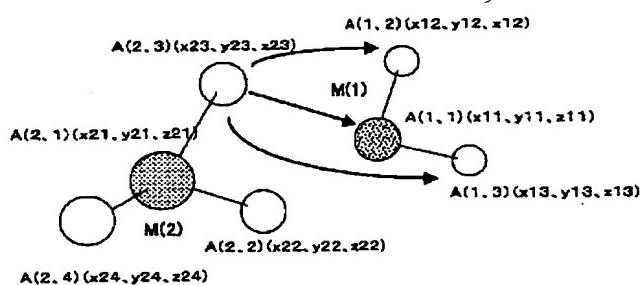
【図1】



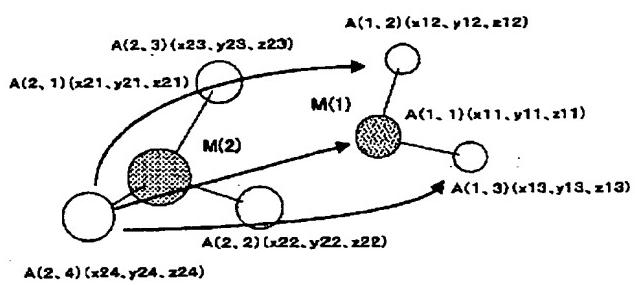
【図2】



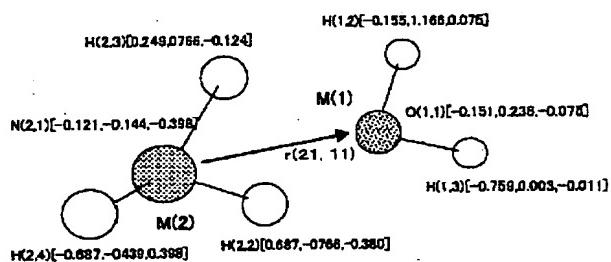
【図3】



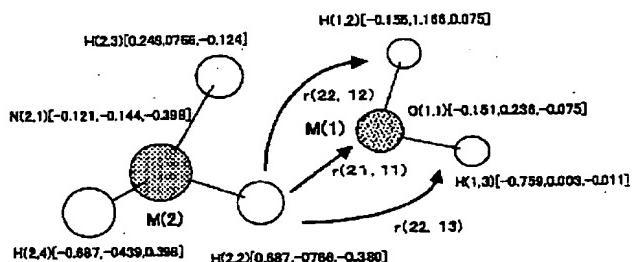
【図4】



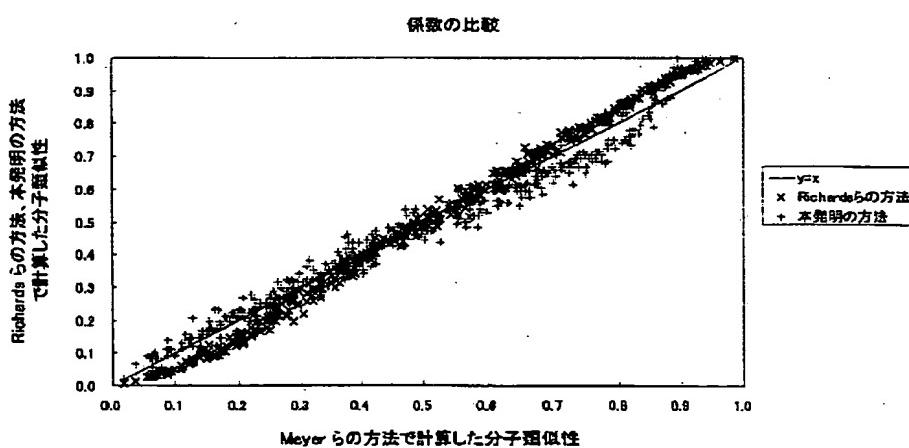
【図6】



【図7】



【図8】



【図5】

